Entropy-partition of Complex Systems and Emergence of Human Brain's Consciousness

Xi Guangcheng

（Institute of Automation, Chinese Academy of Science, Beijing, 100080, P.R. China）
Guangcheng.xi@ia.ac.cn

Abstract. Consciousness is discussed from viewpoint of theory of Entropy—partition of complex system. Human brain's system self-organizably and adaptively implements partition 、 aggregation and integration, and consciousness emerges.

1. Introduction

In article of Giulio Tononi and G.M. Edelman[1], studying of consciousness is concentrated on description of kinds of neural processes which can account for key properties of conscious experience. They have emphasized two key properties of consciousness: conscious experience is integrated (each conscious scene is unified) and at the same time it is highly differentiated (within a short time, one can experience any of a huge number of different conscious states), and have proposed notions so-called functional clustering, neural complexity and dynamical core hypothesis. Those researching methods and train of thought to problem of consciousness agree to a certain extent, with theory and method of Entropy-partition of Complex System I proposed 23 years ago [2]. At that time related specialists (appraising group of achievements in scientific research) ever have pointed out: theory of Entropy-partition (Aggregation) of complex system not only has been applied to Ecologico-Economical regionalization [3] of Ecologico-Economico-social complex system but also will certainly be applied to researching of human Brain's neural system. Therefore having read the article of G. Tononi and G.M. Edelman, I have much sensation seem to have met before, and excited feeling arises spontaneously. Though such, we strongly have felt it still necessary to discuss problem of consciousness from viewpoint of theory of entropy-partition (aggregation) of complex system.

2. Entropy-partition (aggregation) of neural system X

Suppose neural system $X = (X_1, X_2, \cdots, X_a, \cdots, X_p)^T$ is consisted of p neuron, $p \in N$ (N set of natural number), where $X_a = (X_{a_i})$, $a = 1, 2, \cdots, p$; $i = 1, 2, \cdots, q$. Let $C_a (a = 1, \cdots, p)$ be set of classification of X_a , $C_{a_i} = i$ be i-th element of C_a , then we have $C_a = \{1, 2, \cdots, i \cdots, k\}$, $k \leq q$, and let n_i be quantity for X_a belong to i-th class, then entropy of X_a is defined as

$$H(X_a) = -\sum_{i=1}^{k} n_i / q \log n_i / q \tag{1}$$

joint entropy of X_a, X_b is similarly defined as

$$H(X_a \bigcup X_b) = -\sum_i \sum_j n_{ij} / q \log n_{ij} / q \tag{2}$$

where n_{ij} is quantity for X_a belong to i-th class of C_a simultaneously X_b belong to j-th class of C_b . For the convenience of application, expressions (1) and (2) can respectively be represented as

$$H(X_a) = \log q - \frac{1}{q} \sum_{i=1}^{k} n_i \log n_i \tag{1'}$$

$$H(X_a \bigcup X_b) = \log q - \frac{1}{q} \sum_i \sum_j n_{ij} \log n_{ij} \tag{2'}$$

Having had above-mentioned definition of entropy, in what follows, correlative measure by which statistical dependence between the X_a and the X_b is denoted is defined by their mutual information.

Definition 1. Suppose $X_a \bigcap X_b = \phi$, then entropy

$$H(X_a, X_b) = H(X_a) + H(X_b) - H(X_a \bigcup X_b) \tag{3}$$

is called correlative measure $\mu(X_a, X_b)$ between the X_a and the X_b .

Definition 2. Suppose $X_a \bigcap X_b = \phi$ for arbitrary $a, b (a \neq b)$, then

$$\mu(X_1, X_2, \cdots, X_p) \overset{\Delta}{=} \sum_{a=1}^{p} H(X_a) - H\left(\sum_{a=1}^{p} X_a\right) \tag{4}$$

is called correlative measure among X_1 , X_2 , $\cdots$, and X_p .

Definition 2'. Suppose system X be partitioned into m subsystems $s_1, s_2, \cdots, s_m$, for arbitrary i, j $(i \neq j)$, $s_i \bigcap s_j = \phi$, $X = \sum_{i=1}^{m} s_i$, then

$$\mu(s_1, s_2, \cdots, s_m) \stackrel{\Delta}{=} \sum_{i=1}^{m} H(s_i) - H\left(\sum_{i=1}^{m} s_i\right) \tag{5}$$

is called correlative measure among $s_1, s_2, \cdots, s_m$.

Let us consider nonempty finite set X and set-family $E(X)$ consisted of its subset, P_e a set-function defined on $E(X)$ with properties

(i) $P_e(A) \geq 0$, $\forall A \in E(X)$,

(ii) $P_e(\phi) = 0$

Definition 3. If for arbitrary nonempty finite set $S_i \in E(X)$, $S_j \in E(X)$, $i \neq j$, $S_i \bigcap S_j = \phi$, have

$$P_e(S_i \bigcup S_j) \geq P_e(S_i) + P_e(S_j) \tag{6}$$

then, set-function P_e satisfied conditions (i),(ii) is called superadditive.

Although finite superadditivities of mutual information or the so-called measure of cohesion of the components of the set of entities is well known result, it is not still unnecessary to write down our proof about finite superadditivity of the correlative measure among some subsystems of the system X which is given.

Theorem[2]. Correlative measure $\mu(s_1, s_2, \cdots, s_m)$ is finitely superadditive, countably superadditive and unique.

Proof. ***Finite superadditivity.***

Suppose system X is partitioned into m subsystems s_1, s_2, ... , s_m, and that $s_i \in R, s_i \neq \varnothing; s_j \in R, s_j \neq \varnothing; s_i \bigcap s_j = \varnothing$, $X = \sum_{j=1}^{m} s_j \in R$, for any $i, j(i \neq j)$, where R is a algebra on set X which is given. By definition of the correlative measure, we have

$$\mu(S) = \mu(\sum_{j=1}^{m} s_j) = \mu(s_1, s_2, ..., s_m) = \mu(X_1, X_2, ...X_p) = \sum_{i=1}^{p} H(X_i) - H(\sum_{i=1}^{p} X_i) \tag{7}$$

$$\sum_{s_j \in S} \mu(s_j) = \sum_{s_j \in S} (\sum_{X_j \in s_j} H(X_j) - H(\sum_{X_j \in s_j} X_j))$$

$$= \sum_{i=1}^{p} H(X_i) - \sum_{s_j \in S} H(s_j) \qquad (8)$$

$$= \sum_{i=1}^{p} H(X_i) - \sum_{j=1}^{m} H(s_j)$$

Subtracting (8) from (7) leads to

$$\mu(\sum_{j=1}^{m} s_j) - \sum_{j=1}^{m} \mu(s_j) = \sum_{j=1}^{m} H(s_j) - H(\sum_{i=1}^{p} X_i)$$

$$= \sum_{j=1}^{m} H(s_j) - H(\sum_{j=1}^{m} s_j) \geq 0 \qquad (9)$$

Countable superadditivity and uniqueness:

Suppose system X is partitioned into a sequence of subsystems $s_1, s_2, ..., s_n...$, $s_n \in R -$ algebra on X , $s_n \in R,$ $s_n \neq \varnothing;$ $s_m \in R, s_m \neq \varnothing; s_n \bigcap s_m = \varnothing$, $X = \sum_{n=1}^{\infty} s_n \in R$, for any $n, m(n \neq m)$.

Let $E_k = \sum_{n=1}^{k} s_n$, obviously, we have

$$E_1 \subset E_2 \subset \cdots \subset E_k \cdots, E_k \in R, \lim_{k \to \infty} E_k = \sum_{n=1}^{\infty} s_n \in R \qquad (10)$$

From nonnegativity and monotonicity of μ , we have

$$\mu(E_1) \leq \mu(E_2) \leq \cdots \leq \mu(E_k) \leq \cdots \leq \mu(\sum_{n=1}^{\infty} s_n) \qquad (11)$$

If $\lim_{k \to \infty} \mu(E_k) \to \infty$, from monotonicity of μ , then inevitably we have

$\mu(\sum_{n=1}^{\infty} s_n) = \infty$. Hence by using finite superadditivity, we have

$$\mu(\sum_{n=1}^{\infty} s_n) = \infty = \lim_{k \to \infty} \mu(E_k) = \lim_{k \to \infty} \mu(\sum_{n=1}^{k} s_n) \geq \lim_{k \to \infty} \sum_{n=1}^{k} \mu(s_n) = \sum_{n=1}^{\infty} \mu(s_n)$$

If $\lim_{k \to \infty} \mu(E_k) = c < \infty$, above-mentioned process of proof is understood easily. And either of the

$\lim_{k \to \infty} \mu(E_k) = c < \infty$ and $\lim_{k \to \infty} \mu(E_k) = \infty$ makes uniqueness of the limit to hold still, i.e.

$\lim_{k \to \infty} \mu(E_k) = \lim_{k \to \infty} \mu(\sum_{n=1}^{k} s_n)$ is unique. Up to this point, proof of the theorem is completed.

From above-mentioned property of correlative measure, criterion by which lowest requirement of ideal partition is presented should be

(i) $\mu(s_i) > \mu(s_i, s_j)$, $\mu(s_j) > \mu(s_i, s_j)$

(ii) $\mu(s_i) > \mu(s_{i_1}) + \mu(s_{i_2})$, $\mu(s_j) > \mu(s_{j_1}) + \mu(s_{j_2})$.

for any $i, j(i \neq j)$, $s_i \bigcap s_j = \phi$. Here s_{i_l}, s_{j_l} $(l = 1,2)$ denotes partition of s_i and s_j, respectively. (i) denotes that correlative measure of any subsystem itself which is obtained by partition is larger than correlative measure between it and any subsystem. (ii) denotes that correlative measure of any subsystem which is obtained by partition possesses strictly superadditivity.

When number of characteristic variable of system X is very large, comprehensive observation for data is impossible. Even though the observation is possible, as obtaining complete data spends very long time and for other reason such that obtained data loses its available value. At this time, statistical method and theory can be applied to obtaining data and to analyzing problem.

Suppose $\boldsymbol{x}_a = (x_{a_1}, x_{a_2}, \cdots, x_{a_N})$ be index of quantity of characteristic variable X_a of the complex system X , data obtained by mean of random sampling is $\overline{\boldsymbol{x}}_a = (x_{a_1}, x_{a_2}, \cdots, x_{a_q})$, $\overline{\boldsymbol{x}}_a$ be a random sample from population X_a . Obviously any x_{a_i} necessarily is equal to some $x_{a_{\theta_i}}$, here θ_i $(i = 1,2,\cdots,q)$ is number of i-th individual of $\boldsymbol{x}_a$. For any θ_i , we have probability

$$p\left\{x_{a_1} = x_{a_{\theta_1}}, x_{a_2} = x_{a_{\theta_2}}, \cdots, x_{a_q} = x_{a_{\theta_q}}\right\} = (N-q)! / N! \tag{12}$$

Corresponding to expressions (1′) and (2′), we have

$$H(\overline{\boldsymbol{x}}_a) = \frac{(n-q)!}{n!}\left[\log(n!q) - \log(n-q)! - \frac{1}{q}\sum_i n_i \log n_i\right] \tag{13}$$

$$H(\overline{\boldsymbol{x}}_a \bigcup \overline{\boldsymbol{x}}_b) = \frac{(N-q)!}{N!}\left[\log(N!q) - \log(N-q)! - \frac{1}{q}\sum_i \sum_j n_{ij} \log n_{ij}\right] \tag{14}$$

At this time, correlative measure of the system X is defined by way similar to the above-mentioned method. Actually, in order to obtain the partition of the complex system X , we often introduce coefficient of correlative measure $\mu_{ab} = \mu(X_a, X_b) / H(X_b)$ $(\mu_{ij} = \mu(s_i, s_j) / H(s_j)$, obviously, $\mu(a,b)(\mu_{ij})$ is between 0 and 1. Having computed μ_{ab} for all a, b , those X_a ,

X_b whose correlative measure is larger than other $\mu(a,l)$, $l \neq b$ or $\mu(r,b)$, $r \neq a$, are combined so as to accord with the preceding criterion（$\dot{1}$），（$\dot{2}$）. Hence some corresponding subsystems s_i, $i = 1, \cdots, m$, i.e. some neural functional units are obtained.

From neurophysiological studies, elementary unit representing information in human brain is mutually coordinated clique of neuron; by combining mutual function, neurons with different properties carry out correlative activity, form various dynamic neural networks which correspond to various function of information.

Above-mentioned partition or aggregation of neurons is carried out in self-organizable and adaptive form; the partition or aggregation of neurons makes human brain emerge intelligence.

3. Emergence of consciousness

Suppose system X self-organizably and adaptively is partitioned into m subsystem. Correspondingly, there are m subspace $(\Omega_i, \Im_i, P_i)$, $i = 1, \cdots, m$, of probability-space $(\Omega, \Im, P)$, where Ω is space of configuration of system X. Let $(\Omega, \Im, P)$ be Radon measurable space, then for any $\Im_i$ there exists regular conditional probability $P(\cdot \mid \Im_i)$, obviously, $P(\cdot \mid \Im_i) = P(\cdot \mid \Im(s_i)) = P(\cdot \mid s_i) \overset{\Delta}{=} \mu(\cdot \mid s_i)$. Hence, for any $A \in \Im$

$$\mu(A) = \int_{X|s} \mu(\cdot \mid s_i) d\mu(s_i) \tag{15}$$

and for any random variable η

$$E\eta = \int_{X|s} E(\eta \mid s_i) d\mu(s_i) \tag{16}$$

where $X \mid s$ is quotient space.

Let us define following two functions,

$$f_\varepsilon^{(1)}(t) = \begin{cases} \dfrac{\mu(A)}{\varepsilon}, & 0 \leq t \leq \varepsilon \\ 0, & \varepsilon < t \end{cases} \tag{17}$$

$$f_\varepsilon^{(2)}(t) = \begin{cases} \dfrac{E\eta}{\varepsilon}, & 0 \leq t \leq \varepsilon \\ 0, & \varepsilon < t \end{cases} \tag{18}$$

where ε is arbitrary positive number, t is time.

To be this, we eventually have conscious experience, that is $\mu(A)$ or $E(\eta)$ under some situations or $\lim_{\varepsilon \downarrow 0} f_\varepsilon^{(i)}(t)$, $i = 1,2$, under other situations.

When A or η comes from outside of X, or when A or η is caused by "gene" or "experience" which is in inside of X, all will emerge consciousness, and is emerged by system X self-organizably and adaptively to be partitioned (aggregated), then to be integrated in hundreds of milliseconds. $\mu(A)$ or $E\eta$ is the integrated process.

If there exists sequences of mutual information between any appointed two neurons and between any appointed two subsystems, and the sequences of the mutual information have nice asymptotical properties, that is, if there exists ergodic superadditive process for the sequence of mutual information in neural system X, then the system can implement self-organization and self-adaptation.

Remarks

(1) That the neural system X is self-organizably and adaptively partitioned into some subsystems is implemented by means of the correlative measure, i.e. mutual information, which does not touch upon change of location of space of neurons or subsystems. This is a "Internet" in the neural System X. The some subsystems form large-scale hierarchical intelligent system. On the hierarchical intelligent system, basic principle of hierarchical intelligent system IPDI (Increasing precision with decreasing intelligence) holds [4,5]. The layer with high intelligence is dynamic core of the neural system X

(2) Consciousness is some memory.

(3) Unconsciousness is also consciousness.

(4) Consciousness is conscious of consciousness.

(5) Nearly 3000 years ago, Chinese ancient learned men said consciousness is that Saint hears at silence, looks at immateriality, and firmly believed memory all is in brain.

(6) Oneself-consciousness can only be experienced; it forever cannot be captured by oneself.

References

1. Giulio Tonoi and Gerald M. Edelman, Consciousness and Complexity, Science 828, 1846-1851 (1998)
2. Xiguangcheng, Entropy-method of Partition of Complex System, ACTA AUTOMATICA SINICA (in Chinese) 13(3), 216-220, (1987).
3. Xi Guangcheng, The Entropy-method of Eologico-Economical Regionalization, ACTA AUTOMATICA SINICA (in Chinese) 16(2), 170-173, (1990)

8   **Xi** Guangcheng

4. Xiguangcheng, Intelligent Control with Relative Entropy Minimizing, Control Theory and Applications (in Chinese) 16(1), 27-31,(1999)
5. Saridis G N, Analytic formulation of the principle of increasing precision with decreasing intelligence for intelligent machine, Automatica 25(3), 461-467 (1989).